DOI:
10.1136/jnnp-2019-322614

Document Version
Peer reviewed version

A prospective, controlled study of non-motor effects of subthalamic stimulation in Parkinson's disease – results at the 36-month follow-up

Stefanie T. Jost, MSc,[1] Anna Sauerbier, MD/PhD,[1,2] Veerle Visser-Vandewalle, MD/PhD,[3] Keyoumars Ashkan, MD/PhD,[2] Monty Silverdale, MD/PhD,[4] Julian Evans, MD/PhD,[4] Philipp A. Loehrer, MD,[5] Alexandra Rizos, MSc,[2] Jan Niklas Petry-Schmelzer, MD,[1] Paul Reker, MD,[1] Gereon R. Fink, MD/PhD,[1] Jeremy Franklin, PhD,[6] Michael Samuel, MD/PhD,[2] Alfons Schnitzler, MD/PhD,[7] Michael T. Barbe, MD,[1] Angelo Antonini, MD/PhD,[8,9] Pablo Martinez-Martin, MD/PhD,[10] Lars Timmermann, MD,[5] K. Ray Chaudhuri, MD/PhD,[2,11] Haidar S. Dafsari, MD[1]

On behalf of EUROPAR and the International Parkinson and Movement Disorders Society Non-Motor Parkinson's Disease Study Group.

[1]University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Cologne, Germany

[2]Parkinson Foundation International Centre of Excellence, King's College Hospital, London, United Kingdom

[3]University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Stereotaxy and Functional Neurosurgery, Cologne, Germany

[4]Department of Neurology and Neurosurgery, Salford Royal NHS Foundation Trust, Manchester Academic Health Science Centre, University of Manchester, Greater Manchester, United Kingdom

[5]Department of Neurology, University Hospital Giessen and Marburg, Campus Marburg, Marburg, Germany

[6]University of Cologne, Institute of Medical Statistics and Computational Biology (IMSB), Cologne, Germany

[7]Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[8]Parkinson and Movement Disorders Unit, IRCCS Hospital San Camillo, Venice, Italy

[9]Department of Neurosciences (DNS), Padova University, Padova, Italy

[10]Center for Networked Biomedical Research in Neurodegenerative Diseases (CIBERNED), Carlos III Institute of Health, Madrid, Spain

[11]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

Search terms:
Subthalamic nucleus, deep brain stimulation, non-motor symptoms, nonmotor symptoms, quality of life

Submission Type: Article

Number of Tables and Figures: 6

Word Count abstract: 246/250

Word Count of Paper: 3747

Number of References: 43

Corresponding Author:

Haidar S. Dafsari

Department of Neurology

University Hospital Cologne

Kerpener Str. 62

50937 Cologne, Germany

Tel: +49 221 478-37445

email: haidar.dafsari@uk-koeln.de

Statistical analysis conducted by Stefanie T. Jost, University of Cologne, Germany.

Author email addresses
Stefanie T. Jost – stefanie.jost@uk-koeln.de, Anna Sauerbier – anna.sauerbier@uk-koeln.de, Veerle Visser-Vandewalle – veerle.visser-vandewalle@uk-koeln.de, Keyoumars Ashkan – k.ashkan@nhs.net, Monty Silverdale – monty.silverdale@srft.nhs.uk, Julian Evans – julian.evans@srft.nhs.uk, Philipp A. Loehrer – philipp.loehrer@uk-gm.de, Alexandra Rizos – alex.rizos@googlemail.com, Jan Niklas Petry-Schmelzer – jan.petry-schmelzer@uk-koeln.de, Paul Reker – paul.reker@uk-koeln.de, Gereon R. Fink – gereon.fink@uk-koeln.de, Jeremy Franklin – jeremy.franklin@uni-koeln.de, Michael Samuel – m.samuel@nhs.net, Alfons Schnitzler – schnitza@med.uni-duesseldorf.de, Michael T. Barbe – michael.barbe@uk-koeln.de, Angelo Antonini – angelo3000@yahoo.com, Pablo Martinez-Martin – pmm650@hotmail.com, Lars Timmermann – lars.timmermann@uk-gm.de, K. Ray Chaudhuri – chaudhuriray@hotmail.com, Haidar S. Dafsari – haidar.dafsari@uk-koeln.de

Abbreviations: LEDD = levodopa equivalent daily dose; MED = standard-of-care medical therapy; NMS = non-motor symptoms; NMSS= NMSScale; PD = Parkinson's Disease; PDQ-8 SI = 8-item PD Questionnaire Summary Index; QoL = Quality of Life; SCOPA = Scales for Outcomes in PD; STN-DBS = subthalamic nucleus deep brain stimulation

Abstract

Objective: To examine 36-month effects of bilateral subthalamic stimulation (STN-DBS) on non-motor symptoms (NMS) compared to standard-of-care medical therapy (MED) in Parkinson's disease (PD).

Methods: Here we report the 36-month follow-up of a prospective, observational, controlled, international multicenter study of the NILS cohort. Assessments included NMSScale (NMSS), PDQuestionnaire-8 (PDQ-8), Scales for Outcomes in PD (SCOPA)-motor examination, -activities of daily living, and -complications, and levodopa equivalent daily dose (LEDD). Propensity score matching resulted in a pseudo-randomized sub-cohort balancing baseline demographic and clinical characteristics between the STN-DBS and MED group. Within-group longitudinal outcome changes were analyzed using Wilcoxon signed-rank and between-group differences of change scores with Mann-Whitney U test. Strength of clinical responses was quantified with Cohen's effect size. Additionally, bivariate correlations of change scores were explored.

Results: Propensity score matching applied on the cohort of 151 patients (STN-DBS n=67, MED n=84) resulted in a well-balanced sub-cohort including 38 patients per group. After 36 months, STN-DBS significantly improved NMSS, PDQ-8, SCOPA-motor examination and –complications and reduced LEDD. Significant between-group differences, all favoring STN-DBS, were found for NMSS, SCOPA-motor complications, LEDD (large effects), motor examination, and PDQ-8 (moderate effects). Furthermore, significant differences were found for the sleep/fatigue, urinary (large effects), and miscellaneous NMSS domains (moderate effects). NMSS total and PDQ-8 change scores correlated significantly.

Conclusions: This study provides Class IIb evidence for beneficial effects of STN-DBS on NMS at 36-month follow-up which also correlated with quality of life improvements. This highlights the importance of NMS for DBS outcomes assessments.

1. INTRODUCTION

Deep brain stimulation (DBS) of the subthalamic nucleus (STN) improves quality of life (QoL), motor, and non-motor symptoms (NMS) in patients with Parkinson's disease (PD).[1-3] NMS are a key clinical aspect of PD and have been identified as the most important factor in determining QoL in patients with PD.[4] Beneficial non-motor effects in patients undergoing STN-DBS have been observed in clinician-rated as well as laboratory-based investigations of a wide range of NMS, including polysomnography for sleep, urodynamic examinations for urinary symptoms, sniffing sticks for olfaction, as well as sensory signs such as thermal detection thresholds and pain thresholds.[5 6]

Benefits of bilateral STN-DBS in patients with PD on a wide range of NMS have been reported for a follow-up period up to two years.[3] It is unclear, if beneficial non-motor effects of STN-DBS are sustained beyond 24 months. Non-motor effects of STN-DBS have been compared to apomorphine and levodopa infusion.[7] However, little is known about the progression of non-motor symptoms in patients undergoing standard-of-care medical treatment (MED). To our knowledge, a controlled study of non-motor effects of STN-DBS and MED has not been conducted yet. We hypothesized that beneficial motor and non-motor effects of STN-DBS can be observed in a controlled study at 36-month follow-up after surgery.

2. METHODS

2.1 Study design and ethical approval

Here we report the 36-month follow-up of an ongoing, prospective, observational, controlled, international multicenter study (Class IIb evidence). Patients were recruited between 03/2011 and 10/2015 as part of the DBS and medication arms of the NILS study[7] which is a comprehensive study with non-motor profiling of PD as the primary outcome measure addressing the progression of NMS and treatment responses to medication and advanced

treatments. Medical ethics committees of the participating centers approved the study protocol (master votes for Germany: Cologne, study number: 12-145, German Clinical Trials Register: DRKS00006735, and for the United Kingdom: National Research Ethics Service South East London REC 3, 10/H0808/141; NIHR portfolio, number: 10084). The study was carried out in accordance with the declaration of Helsinki. All patients gave written informed consent prior to any study procedures.

2.2 Participants

PD diagnosis was based on the UK Brain Bank criteria.[8] As per clinical routine, patients in the DBS group were screened for DBS surgery according to the international guidelines, as described in previous publications by our group.[7 9] Patients required a sufficient levodopa responsiveness (>30%) as assessed by the Unified PD Rating Scale-motor examination (UPDRS-III). Indications for DBS were based on multi-disciplinary assessments by movement disorders specialists, stereotactic neurosurgeons, neuropsychologists, psychiatrists, and when necessary, speech therapists and physiotherapists. Patients with clinically relevant psychiatric diseases or neuropsychological impairment (Mini-Mental State Examination scores<25) were excluded from this analysis. To ensure comparability between the STN-DBS and MED group, we included only patients in the MED group, who had advanced PD with dyskinesia, ON/OFF fluctuations or medication-refractory tremor. According to the multi-disciplinary assessments, patients in the MED group were considered candidates for STN-DBS but at that time preferred conservative non-surgical therapy according to published standard-of-care recommendations.[10] Patients' informed decisions were shaped by a number of parameters, such as patients' age, disease duration, dopaminergic medication requirements, and motor and non-motor symptom profiles. During the course of the study, patients in the MED group could undergo DBS at any time and patients in the DBS group could switch off neurostimulation at any time.

2.3 Clinical assessment and outcome parameters

Patients in both groups were assessed in the on-medication state (MedON) at baseline and 36-month follow-up. The STN-DBS group was assessed in the medication and stimulation ON state (MedON/-StimON) at follow-up.

The principal outcome measure was the Non-motor Symptom Scale (NMSS).[11] The 30-item NMSS evaluates nine dimensions of NMS in PD: cardiovascular, sleep/fatigue, mood/apathy, perceptual problems/hallucinations, attention/memory, gastrointestinal symptoms, urinary symptoms, sexual function, and miscellaneous symptoms including items for pain, inability to smell/taste, weight change, and excessive sweating. The total score ranges from 0 (no impairment) to 360 (maximum impairment).

Furthermore, assessments included the following secondary outcome measures:

QoL was quantified by means of the PD Questionnaire (PDQ)-8, a short-form of the PDQ-39[12] covering eight dimensions contributing to QoL. The scale is recommended for QoL assessments by the Movement Disorders Society Scales Committee and has been commonly used in PD and in patients undergoing DBS.[13 14] The PDQ-8 is reported as Summary Index (PDQ-8 SI) ranging from 0 (no impairment) to 100 (maximum impairment).

The Scales for Outcomes in PD (SCOPA) was used to assess motor examination, activities of daily living, and motor complications. The scale is a well-established, validated short version of the Unified Parkinson's Disease Rating Scale[15], highly correlates with the corresponding UPDRS parts[16] and its assessment time is approximately four time shorter than in the MDS-Unified Parkinson's disease rating scale.[15 17] SCOPA subscales range from 0 (no impairment) to 42 (motor examination), 21 (activities of daily living), and 12 (motor complications).[3]

The levodopa equivalent daily dose (LEDD) was computed according to the method by Tomlinson et al.[18]

2.4 Statistical analysis

Propensity score matching was used to create a pseudo-randomized[19] sub-cohort balancing the following baseline demographic and disease parameters between the STN-DBS and MED group: age at baseline, disease duration since diagnosis, LEDD and SCOPA total score. The matching was conducted with Propensity Score Matching for SPSS (version 3.04)[20] and SPSS 25.0 (IBM Corporation). We implemented a 1:1 ratio nearest-neighbor matching algorithm with a 0.25 caliper without replacement and conducted balance diagnostics based on Cohen's effect size $|d|<0.25$.[21]

The assumption of normality was assessed with the Shapiro-Wilk test. Differences of baseline characteristics between the two treatment groups for dichotomous variables were analyzed using the Chi^2-test and for continuous variables using Mann-Whitney U tests or unpaired t-tests, when parametric criteria were fulfilled. To determine outcome changes from baseline to follow-up within each group, Wilcoxon signed-rank or paired samples t-tests were calculated. Mann-Whitney U tests of change scores between STN-DBS and MED groups (mean $test_{baseline}$ − mean $test_{follow\text{-}up}$) were conducted to analyze differences of outcome parameter changes. Multiple comparisons resulting from the use of 7 scales were corrected with the Benjamini-Hochberg procedure to account for multiple testing. All p values presented are two-sided and were adjusted to give a global family-wise significance threshold $p<0.05$. To determine clinical relevance of the responses, we calculated relative change $[(mean\ test_{baseline} - mean\ test_{follow-up}) / mean\ test_{baseline}]$ and Cohen's d effect size with a method by Morris for pretest-posttest-control group designs.[22] Furthermore, the relationship between NMSS and PDQ-8 change scores was explored using Spearman correlation.

3. RESULTS

Of 317 patients screened, 151 were treated with STN-DBS or MED and met inclusion criteria (figure 1). Of 151 patients in the final sample (100 male) with a mean age of 63.8 ±9.8 years,

67 patients underwent STN-DBS and 84 MED. The mean time to follow-up was 3.1 years. None of the patients in the MED group underwent STN-DBS during the 36-month follow-up period.

3.1 Baseline characteristics in the original and matched cohort

In the original cohort, patients in the STN-DBS group were significantly younger (61.3 ±8.5 years vs. 65.8 ±10.4 years) with longer duration of PD (10.8 ±4.9 years vs. 7.5 ±5.0 years) and had greater dopaminergic medication requirements (LEDD 1199.1 ±587.5 mg/day vs. 778.1 ±407.1 mg/day) and more severe impairment in most of the clinical scales (table 1).

Propensity score matching resulted in a sub-cohort of 76 patients (38 patients in each group). Balance diagnostics indicated a good matching between the two groups, i.e. no significant differences were found for the main demographic and clinical outcome parameters. Differences in NMSS domain scores were not significant except for sleep/fatigue. Baseline motor subscores and Mini-Mental State Examination scores of the matched cohort are available in the Supplementary table e-1.

The results reported in this manuscript relate to the matched cohort. Additionally, outcome changes in the original cohort are reported in the Supplementary table e-2.

3.2 Clinical Outcomes at Baseline and 36-month follow-up

The time of follow-up assessment did not differ significantly between the groups (STN-DBS=3.1 ±0.24 years, MED=3.1 ±0.47 years; p=0.580). Longitudinal within-group changes and between-group differences are reported in table 2. The STN-DBS group significantly improved in NMSS total score, PDQ-8 SI, SCOPA-motor examination, and -complications. No significant longitudinal change was found for SCOPA-activities of daily living. In the MED group, outcome parameters did not change significantly. As expected, LEDD was significantly reduced in the STN-DBS group, while it remained stable in the MED group.

Post-hoc exploratory analyses of NMSS domains (see table 2 and figure 2) resulted in significant beneficial effects of STN-DBS for the sleep/fatigue, urinary, sexual function and miscellaneous domains. The latter was driven by pain (baseline, 2.2 ±3.3; follow-up, 0.6 ±1.5, p=0.008), inability to smell/taste (baseline, 3.5 ±4.0; follow-up, 1.6 ±3.2, p=0.008) and weight changes (baseline, 1.8 ±3.1; follow-up, 0.6 ±1.5, p=0.025). A significant worsening in the STN-DBS group was found in the cardiovascular domain. In the MED group, the domains sleep/fatigue and urinary worsened significantly.

Favoring STN-DBS, significant differences between change scores of STN-DBS and MED groups were found for the NMSS total score and the sleep/fatigue, urinary, and miscellaneous domains. The latter was driven by the inability to smell/taste and pain. Furthermore, favoring STN-DBS, significant differences were observed for PDQ-8 SI, LEDD, SCOPA-motor examination, and -complications.

Table 3 shows change scores and relative changes from baseline to 36-month follow-up for both groups and Cohen's d effect size for the differences in change scores between the two treatment groups.

In the DBS group, we observed 6 Serious Adverse Events in 5 patients (skin perforation over battery site in two patients, disturbed wound healing, dopamine agonist withdrawal syndrome, suicide attempt, mania) which all resolved without major sequelae and 68 neurological Adverse Events in 40 patients and 25 psychiatric Adverse Events in 20 patients.

3.3 Spearman correlation analyses

There was a significant correlation of 'moderate' strength between change scores of NMSS total and PDQ-8 SI (see table 4). Explorative correlation analyses between change scores in NMSS domains and PDQ-8 SI revealed significant correlations for the domains sleep/fatigue ('moderate'), mood/apathy, attention/memory, urinary, and miscellaneous (all 'weak'). The

correlation within the miscellaneous domain was driven by the inability to smell/taste (r_s=0.354, p=0.002) and weight changes (r_s=0.231, p=0.045).

Additionally, Supplementary table e-3 reports Spearman correlations between change scores of PDQ-8 SI and NMSS for the separate treatment groups. In the STN-DBS and the MED group, LEDD was not correlated to PDQ-8 SI or NMSS (all p>0.05). There was a 'weak' correlation between change scores in PDQ-8 SI and SCOPA-motor examination (r_s =0.333, p=0.005), -activities of daily living (r_s=0.367, p=0.002), and -motor complications (r_s=0.316, p=0.006).

4. DISCUSSION

This study reports Class IIb evidence for beneficial effects of STN-DBS on a wide range of NMS in a controlled design at 36-month follow-up. Patients treated with STN-DBS experienced a better outcome of total NMS burden and specific non-motor aspects, such as sleep/fatigue, urinary symptoms, inability to smell/taste, and pain, than patients treated with MED.

Motor symptoms, LEDD and (Serious) Adverse Events

In line with previous studies, STN-DBS resulted in a significant improvement of motor aspects of PD.[2] Confirming results from a study by Weaver et al., we observed an improvement of motor complications 36 months after STN-DBS.[23] In accordance with this study, we observed a significant sustained 30% LEDD reduction at 36-month follow-up. In the MED group, standard-of-care practice[10] stabilized motor examination, activities of daily living, and motor complications over the 36-month course of the study with no significant increase of dopaminergic medication. As expected, changes of medication requirements differed significantly between the two groups. Based on the sum of patient years, this result is well within the range of studies with long-term follow-up periods.[24]

Non-motor Symptoms

In line with a study by Holmberg et al. reporting no significant difference of *cardiovascular* outcomes between the STN-DBS and MED group at 12-month follow-up[25], we observed no significant differences in the NMSS cardiovascular domain outcome at 36-month follow-up. However, when regarding longitudinal changes, we observed a significant worsening of cardiovascular symptoms in patients undergoing STN-DBS. The worsening of cardiovascular symptoms was unlikely to result from potential side effects of dopaminergic medication as these were significantly reduced postoperatively. In the STN-DBS group of the original cohort, cardiovascular symptoms at baseline were higher than in the matched sub-cohort and did not worsen at 36-month follow-up. Further studies are needed to investigate long-term effects of STN-DBS on cardiovascular symptoms in patients with high and low baseline cardiovascular impairments.

This is the first controlled study reporting an improvement of subjective *sleep/fatigue* symptoms. Beneficial effects of STN-DBS on subjective sleep symptoms have been demonstrated in studies with follow-up periods up to three years.[6 7 26 27] Our results are in line with a recent study by Choi et al. that found improvements in sleep dysfunction at three-year follow-up.[26] Our study adds to the evidence on beneficial effects on quality of sleep by demonstrating improvements in a controlled design. In contrast, Lilleeng et al. reported no changes of sleep symptoms but a significant worsening of fatigue.[28] However, this result was limited by the fact that the medication requirements at postoperative follow-ups remained high (LEDD at 1.0–1.5 years: 810 mg and at 6.0–9.0 years: 910 mg) and sleep and fatigue symptoms are common side effects of dopaminergic medication.[29] As described in previous studies[30 31], possible other mechanisms of STN-DBS effects on sleep/fatigue, besides LEDD reduction, are: a direct modulation of basal ganglia-thalamo-cortical loops resulting in improved nocturnal motor symptoms and also a spread of current to regions near the STN, e.g., the pedunculopontine nucleus.

In accordance with previous evidence, we observed no significant changes of *mood/apathy* in the STN-DBS group at 36-month follow-up.[23] However, this finding needs to be confirmed, as improved depression at 3-year follow-up after STN-DBS has been reported in Class IV evidence.[32] Future studies addressing depression and its interplay with other neuropsychiatric symptoms, such as anxiety, hypomania, and alexithymia with longer follow-up are needed.[33]

Perceptual problems/hallucinations remained unchanged from baseline to 36-month follow-up in the present study. In one of the few available studies on this issue, Yoshida et al. reported improved hallucinations in STN-DBS at 6-month follow-up.[34] Future studies are needed to investigate the relationship of hallucination outcome and dopaminergic medication, psychotropic co-medication and the co-dependency with other neuropsychiatric aspects of PD.

In our study, we observed no significant changes in *attention/memory* from baseline to 36-month follow-up and no inter-group differences between STN-DBS and MED. This is in line with a study by Funkiewiez et al. that found no significant change of global cognitive functions and attention and memory subscales 3 years after surgery.[32] This is also in line with a study by Zangaglia et al. with multiple follow-up visits up to 3 years which reported no significant worsening in memory tasks after STN-DBS, whereas verbal fluency performance deteriorated in the STN-DBS group and logical executive function tasks were only impaired transiently at 12-month follow-up. [35]

In our cohort we found no significant change of *gastrointestinal symptoms* at 36-month follow-up. A study by Zibetti et al. reported a significantly lower prevalence of constipation at 24-month follow-up.[36] However, the authors did not employ validated scales and retrospectively extracted dichotomized data on the presence of constipation from patient charts.

To our knowledge, this is the first study to report beneficial 36-month effects of STN-DBS on *urinary symptoms* compared to a MED control group. This is in line with previous studies with shorter follow-up periods.[3] Herzog et al. reported ameliorations of bladder dysfunctions along with a modulation of blood flow of the posterior thalamus and the insular cortex.[37] Nonetheless,

long-term effects are unclear as Yamamoto et al. found no significant change in urodynamic parameters three years after surgery, which may have resulted from a relatively small sample size (n=13).[38]

In the present study, *sexual functions* improved significantly in the STN-DBS group. This is in line with a study which reported an improvement of sexual function in 21 male patients undergoing STN-DBS in PD at a 9–12-month follow-up.[39] In contrast, Kurcova et al. found no significant changes of sexual function in four female patients at a 4-month follow-up.[40] Therefore, the effect of STN-DBS on sexual function may depend on demographic parameters, such as sex of the patient.

The present study shows beneficial effects on the *miscellaneous* domain. In our study, STN-DBS significantly improved *pain* at 36-month follow-up and this beneficial effect was also significant in the between-group comparison. These results are consistent with previous reports of beneficial effects of STN-DBS on pain in a 1- to 8-year follow-up after surgery. [6 41] The mechanisms underlying the impact of STN-DBS on pain are not fully understood. Sensory gating has been discussed as a possible mechanism.[31] To our knowledge, the present study is also the first to report beneficial effects of STN-DBS on subjective *olfactory* function at 36-month follow-up. This is in line with studies showing beneficial immediate effects on odor identification[6] and a PET study by Cury et al. provided evidence for an increased glucose metabolism in the midbrain and right frontal lobe in patients with improved olfactory function following STN-DBS.[42] Our study extends the time frame of beneficial effects to 36 months after STN-DBS. We found no significant difference between the STN-DBS and the MED group for *excessive sweating* and for *weight changes*. In the literature, only few studies have reported significant changes of these non-motor aspects of PD in patients undergoing STN-DBS and for both symptoms the sample sizes of available studies were small and the follow-up periods short.[6] A study by Petry-Schmelzer et al. provides a good overview of possible mechanisms of

action (location-specific and general) which may mediate beneficial non-motor effects of STN-DBS.[31]

Quality of life

Confirming the results of earlier studies, we observed a significant QoL improvement at 36-month follow-up in the within-group analysis of the STN-DBS group.[23] In the MED group, standard-of-care practice stabilized QoL over the 36-month course of the study. In line with previous studies with shorter follow-up periods up to two years,[1] we observed a significantly greater QoL improvement after STN-DBS compared to MED at 36-month follow-up. QoL and NMS total burden correlated significantly in the overall matched cohort as well as in the STN-DBS and MED groups separately. QoL correlated significantly with sleep/fatigue, mood/apathy, attention/memory, urinary, and the miscellaneous NMSS domains. QoL and motor examination were also significantly associated, but the correlation was weaker than with non-motor aspects, which is in line with previous studies and highlights the relative importance of NMS.[43] QoL changes were significantly correlated to mood/apathy and attention/memory changes, indicating that although group level clinical effects *per se* were not significant, these non-motor outcomes were significantly related with postoperative QoL outcome on an individual level.

Limitations

There are several limitations in our study to be considered. While propensity score matching is a well-established tool to precisely match baseline characteristics between two groups, this method can only be applied to parameters assessed clinically. Therefore, this method does not consider potentially relevant parameters, which were not measured, for example impulse control disorders. To account for this limitation, comparisons between the matched groups in this study were only carried out conservatively using independent samples tests. Diagnostic statistics indicated a well-balanced matching for the selected matching parameters, and this also

led to a good balance for most NMSS domains. A notable difference, however, was the mean NMSS sleep/fatigue domain, which was higher at baseline in the STN-DBS group than in the MED group and may have contributed to the greater improvement observed in the STN-DBS group. Conversely, we found low baseline impairment in the STN-DBS group, e.g., for the cardiovascular domain which may have left little room for additional improvements (floor effect). While propensity score matching can, therefore, not replace randomized controlled trials, it can increase causal inference in observational trials and is an important tool in situations in which the randomized controlled design may not be suitable (e.g. in the present study, investigations of long-term effects which would otherwise have resulted in withholding an effective therapy from severely affected patients for 3 years). In this context, one has to acknowledge, that one of the reasons why patients in the MED group might have decided against DBS therapy at that time, could have been their short disease duration (mean 7.5 years in the original cohort) which was even shorter than in the EARLYSTIM study (7.7 years).[1] Furthermore, in the STN-DBS group, the matching resulted in a selection of less severely affected patients with shorter mean disease duration, as there were too few matching partners from the MED group within the defined caliper. This is important, as the observed effects of STN-DBS may be different in patients with very severe PD. Further studies are required to investigate the dependence of STN-DBS outcomes on the levels of baseline impairments and predict long-term non-motor outcomes. We chose a conservative caliper (0.25) and conducted strict balance diagnostics[21] to implement a precise matching between the two groups. A narrower caliper would have resulted in smaller matched cohort sizes. Nonetheless, to our knowledge, this cohort including 76 matched patients was one of the biggest in studies of its kind. The multicenter design of our study may reduce bias caused by single center studies and thus increase external validity. Another limitation was that the medication changes were not determined by an independent external panel as, e.g., in the EARLYSTIM study.[1] However, the standard-of-care clinical procedures of movement disorder specialists in each center were

based on the same criteria[10] and clinical outcomes, such as PDQ-8 SI, SCOPA-motor examination, -activities of daily living, and -motor complications, and NMSS total burden did not worsen significantly in the MED group over the 36-month follow-up period, indicating a successful treatment of these aspects of PD. Furthermore, we analyzed Cohen's d effect size to help the interpretation of our results with regard to the clinical relevance of the observed changes. However, the ideal method for this purpose would have been the use of minimal clinically important changes, which, to our knowledge, has not been published for the NMSS and its domains yet.

Conclusion

Our controlled study with a 36-month follow-up provides Class IIb evidence for a beneficial effect of STN-DBS on NMS total burden and a wide range of NMS, such as sleep/fatigue, urinary symptoms, pain, and olfactory functions. The clinical relevance of these non-motor outcomes is highlighted by their significant correlation with QoL improvements after STN-DBS. Studies comparing QoL, motor and non-motor effects of different treatment options, e.g. for STN-DBS and standard-of-care medical therapy as investigated in this study, will help to provide a basis on personalized medicine to patient's individual PD profiles.

Appendix 1. Authors.

Name	Location	Role	Contribution
Stefanie T. Jost	University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Germany	Author	data analysis, drafting of the manuscript and figures
Anna Sauerbier	Parkinson Foundation International Centre of Excellence, King's College Hospital, London, United Kingdom; University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Germany	Author	data acquisition, critical revision of the manuscript
Veerle Visser-Vandewalle	University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Stereotaxy and Functional Neurosurgery, Germany	Author	surgical intervention, critical revision of the manuscript
Keyoumars Ashkan	Parkinson Foundation International Centre of Excellence, King's College Hospital, London, United Kingdom	Author	surgical intervention, critical revision of the manuscript
Monty Silverdale	Department of Neurology and Neurosurgery, Salford Royal NHS Foundation Trust, Manchester Academic Health Science Centre, University of Manchester, Greater Manchester, United Kingdom	Author	data acquisition, critical revision of the manuscript
Julian Evans	Department of Neurology and Neurosurgery, Salford Royal NHS Foundation Trust, Manchester Academic Health Science Centre, University of Manchester, Greater Manchester, United Kingdom	Author	surgical intervention, critical revision of the manuscript
Philipp A. Loehrer	Department of Neurology, University Hospital Giessen and Marburg, Campus Marburg, Marburg, Germany	Author	data acquisition, critical revision of the manuscript
Alexandra Rizos	Parkinson Foundation International Centre of Excellence, King's College Hospital, London, United Kingdom	Author	data acquisition, critical revision of the manuscript
Jan Niklas Petry-Schmelzer	University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Germany	Author	data acquisition, critical revision of the manuscript
Paul Reker	University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Germany	Author	data acquisition, critical revision of the manuscript
Gereon R. Fink	University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Germany	Author	data acquisition, critical revision of the manuscript
Jeremy Franklin	University of Cologne, Institute of Medical Statistics and Computational Biology (IMSB), Germany.	Author	critical revision of the manuscript

Michael Samuel	Parkinson Foundation International Centre of Excellence, King's College Hospital, London, United Kingdom	Author	critical revision of the manuscript
Alfons Schnitzler	Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany	Author	critical revision of the manuscript
Michael T. Barbe	University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Germany	Author	data acquisition, critical revision of the manuscript
Angelo Antonini	Parkinson and Movement Disorders Unit, IRCCS Hospital San Camillo, Venice, Italy	Author	data acquisition, critical revision of the manuscript
Pablo Martinez-Martin	Center for Networked Biomedical Research in Neurodegenerative Diseases (CIBERNED), Carlos III Institute of Health, Madrid, Spain	Author	study concept and design, critical revision of the manuscript
Lars Timmermann	Department of Neurology, University Hospital Giessen and Marburg, Campus Marburg, Marburg, Germany	Author	data acquisition, critical revision of the manuscript
K. Ray Chaudhuri	Parkinson Foundation International Centre of Excellence, King's College Hospital, London, United Kingdom	Author	study concept and design, data acquisition, critical revision of the manuscript
Haidar S. Dafsari	University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Germany	Author	study concept and design, data acquisition, critical revision of the manuscript

Funding

Stefanie T. Jost has received a travel grant from the German Academic Exchange Service.

Anna Sauerbier reports no financial disclosures.

Veerle Visser-Vandewalle is a member of the advisory boards and reports consultancies for Medtronic, Boston Scientific and St. Jude Medical. She received a grant from SAPIENS Steering Brain Stimulation.

Keyoumars Ashkan has received honoraria for educational meetings, travel and consultancy from Medtronic, St Jude Medical and Boston Scientific.

Monty Silverdale has received honoraria from Bial, Britannia and Medtronic.

Julian Evans reports no financial disclosures.

Philipp A. Loehrer: was funded by the Konrad-Adenauer-Foundation and the Koeln-Fortune Program. He reports travel grants from AbbVie.

Alexandra Rizos has received honorarium from UCB and was supported by a grant from Medtronic.

Jan Niklas Petry-Schmelzer has received travel grants from Boston Scientific.

Paul Reker has received a travel grant from AbbVie.

Gereon R. Fink reports no financial disclosures.

Jeremy Franklin reports no financial disclosures.

Michael Samuel has received honoraria for educational meetings/travel/accommodation from Medtronic, St. Jude Medical, and UCB, grants from Parkinson's UK and Ipsen, and has acted as a consultant for Medtronic and St. Jude Medical.

Alfons Schnitzler reports personal consultancy and lecture fees from Medtronic Inc, Abbott, and Boston Scientific, lecture fees from UCB, Teva Pharma, Bial, and Zambon, and through his institution funding for his research from the German Research Council, the German Ministery of Education and Health.

Michael T. Barbe reports grants from Boston Scientific and Medtronic.

Angelo Antonini reports personal consultancy fees from Zambon, AbbVie, Boehringer Ingelheim, GE, Neuroderm, Biogen, Bial, EVER Neuro Pharma, Therevance, Vectura grants from Chiesi Pharmaceuticals, Lundbeck, Horizon 2020 - PD_Pal Grant 825785, Ministry of

Education University and Research (MIUR) Grant ARS01_01081, owns Patent WO2015110261-A1, owns shares from PD Neurotechnology Limited.

Pablo Martinez-Martin has received honoraria from Editorial Viguera and Movement Disorder Society for lecturing in courses; from AbbVie for speaking in experts' meetings and from AbbVie and Zambon for participating in the Advisory Board of epidemiological studies. License fee payments for the King's Parkinson's Disease Pain Scale, and grants from the International Parkinson and Movement Disorder Society for development and validation of the MDS-Non-Motor Symptoms Scale.

Lars Timmermann reports grants, personal fees and non-financial support from SAPIENS Steering Brain Stimulation, Medtronic, Boston Scientific and St. Jude Medical.

K. Ray Chaudhuri has received funding from Parkinson's UK, NIHR, UCB, and the European Union; he received honoraria from UCB, Abbott, Britannia, US Worldmeds, and Otsuka Pharmaceuticals; and acted as a consultant for AbbVie, UCB, and Britannia.

Haidar S. Dafsari's work was funded by the Prof. Klaus Thiemann Foundation and the Felgenhauer Foundation and has received honoraria by Boston Scientific and Medtronic.

This paper is independent research funded by the German Research Foundation (Grant KFO 219).

References

1. Schuepbach WM, Rau J, Knudsen K, et al. Neurostimulation for Parkinson's disease with early motor complications. *N Engl J Med* 2013;368(7):610-22. doi: 10.1056/NEJMoa1205158.
2. Deuschl G, Schade-Brittinger C, Krack P, et al. A randomized trial of deep-brain stimulation for Parkinson's disease. *N Engl J Med* 2006;355(9):896-908. doi: 10.1056/NEJMoa060281.
3. Dafsari HS, Silverdale M, Strack M, et al. Nonmotor symptoms evolution during 24 months of bilateral subthalamic stimulation in Parkinson's disease. *Mov Disord* 2018;33(3):421-30. doi: 10.1002/mds.27283.
4. Martinez-Martin P, Rodriguez-Blazquez C, Kurtis MM, et al. The impact of non-motor symptoms on health-related quality of life of patients with Parkinson's disease. *Mov Disord* 2011;26(3):399-406. doi: 10.1002/mds.23462.
5. Marques A, Chassin O, Morand D, et al. Central pain modulation after subthalamic nucleus stimulation: A crossover randomized trial. *Neurology* 2013;81(7):633-40. doi: 10.1212/WNL.0b013e3182a08d00.
6. Kurtis MM, Rajah T, Delgado LF, et al. The effect of deep brain stimulation on the non-motor symptoms of Parkinson's disease: a critical review of the current evidence. *NPJ Parkinsons Dis* 2017;3:16024. doi: 10.1038/npjparkd.2016.24.
7. Dafsari HS, Martinez-Martin P, Rizos A, et al. EuroInf 2: Subthalamic stimulation, apomorphine, and levodopa infusion in Parkinson's disease. *Mov Disord* 2019;34(3):353-65.
8. Hughes AJ, Daniel SE, Kilford L, et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992;55(3):181-4.
9. Dafsari HS, Reker P, Stalinski L, et al. Quality of life outcome after subthalamic stimulation in Parkinson's disease depends on age. *Mov Disord* 2018;33(1):99-107. doi: 10.1002/mds.27222.
10. Horstink M, Tolosa E, Bonuccelli U, et al. Review of the therapeutic management of Parkinson's disease. Report of a joint task force of the European Federation of Neurological Societies (EFNS) and the Movement Disorder Society-European Section (MDS-ES). Part II: late (complicated) Parkinson's disease. *Eur J Neurol* 2006;13(11):1186-202. doi: 10.1111/j.1468-1331.2006.01548.x.
11. Chaudhuri KR, Martinez-Martin P, Brown RG, et al. The metric properties of a novel non-motor symptoms scale for Parkinson's disease: Results from an international pilot study. *Mov Disord* 2007;22(13):1901-11.
12. Jenkinson C, Fitzpatrick R, Peto V, et al. The PDQ-8: Development and validation of a short-form Parkinson's Disease Questionnaire. *Psychol Health* 1997;12(6):805-14. doi: Doi 10.1080/08870449708406741.
13. Espay AJ, Vaughan JE, Marras C, et al. Early versus delayed bilateral subthalamic deep brain stimulation for parkinson's disease: a decision analysis. *Mov Disord* 2010;25(10):1456-63. doi: 10.1002/mds.23111.
14. Dafsari HS, Reker P, Silverdale M, et al. Subthalamic Stimulation Improves Quality of Life of Patients Aged 61 Years or Older With Short Duration of Parkinson's Disease. *Neuromodulation* 2018;21(6):532-40. doi: 10.1111/ner.12740.
15. Marinus J, Visser M, Stiggelbout AM, et al. A short scale for the assessment of motor impairments and disabilities in Parkinson's disease: the SPES/SCOPA. *J Neurol Neurosurg Psychiatry* 2004;75(3):388-95.

16. Martinez-Martin P, Benito-Leon J, Burguera JA, et al. The SCOPA-Motor Scale for assessment of Parkinson's disease is a consistent and valid measure. *J Clin Epidemiol* 2005;58(7):674-9. doi: 10.1016/j.jclinepi.2004.09.014.
17. Verbaan D, van Rooden SM, Benit CP, et al. SPES/SCOPA and MDS-UPDRS: formulas for converting scores of two motor scales in Parkinson's disease. *Parkinsonism Relat Disord* 2011;17(8):632-4. doi: 10.1016/j.parkreldis.2011.05.022.
18. Tomlinson CL, Stowe R, Patel S, et al. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov Disord* 2010;25(15):2649-53. doi: 10.1002/mds.23429.
19. Everett BG, Onge JS, Mollborn S. Effects of Minority Status and Perceived Discrimination on Mental Health. *Popul Res Policy Rev* 2016;35(4):445-69. doi: 10.1007/s11113-016-9391-3.
20. Thoemmes F. Propensity Score Matching in SPSS. *arXiv:12016385* 2012.
21. Stuart EA, Rubin DB. Best practices in quasi-experimental designs: matching methods for causal inference. In: Osborne J, ed. Best practices in quantitative social science. Thousand Oaks, CA: Sage 2007.
22. Morris SB. Estimating effect sizes from pretest-posttest-control group designs. *Organ Res Methods* 2008;11(2):364-86. doi: 10.1177/1094428106291059.
23. Weaver FM, Follett KA, Stern M, et al. Randomized trial of deep brain stimulation for Parkinson disease: thirty-six-month outcomes. *Neurology* 2012;79(1):55-65. doi: 10.1212/WNL.0b013e31825dcdc1.
24. Buhmann C, Huckhagel T, Engel K, et al. Adverse events in deep brain stimulation: A retrospective long-term analysis of neurological, psychiatric and other occurrences. *PLoS One* 2017;12(7):e0178984. doi: 10.1371/journal.pone.0178984.
25. Holmberg B, Corneliusson O, Elam M. Bilateral stimulation of nucleus subthalamicus in advanced Parkinson's disease: no effects on, and of, autonomic dysfunction. *Mov Disord* 2005;20(8):976-81. doi: 10.1002/mds.20492.
26. Choi JH, Kim HJ, Lee JY, et al. Long-term effects of bilateral subthalamic nucleus stimulation on sleep in patients with Parkinson's disease. *PLoS One* 2019;14(8):e0221219. doi: 10.1371/journal.pone.0221219.
27. Dafsari HS, Ray-Chaudhuri K, Ashkan K, et al. Beneficial effect of 24-month bilateral subthalamic stimulation on quality of sleep in Parkinson’s disease. *J Neurol* 2020 (in print) doi: 10.1007/s00415-020-09743-1.
28. Lilleeng B, Gjerstad M, Baardsen R, et al. The long-term development of non-motor problems after STN-DBS. *Acta Neurol Scand* 2015;132(4):251-8. doi: 10.1111/ane.12391.
29. Zesiewicz TA, Hauser RA. Sleep attacks and dopamine agonists for Parkinson's disease: what is currently known? *CNS Drugs* 2003;17(8):593-600. doi: 10.2165/00023210-200317080-00004.
30. Dafsari HS, Petry-Schmelzer JN, Ray-Chaudhuri K, et al. Non-motor outcomes of subthalamic stimulation in Parkinson's disease depend on location of active contacts. *Brain Stimul* 2018;11(4):904-12. doi: 10.1016/j.brs.2018.03.009.
31. Petry-Schmelzer JN, Krause M, Dembek TA, et al. Non-motor outcomes depend on location of neurostimulation in Parkinson's disease. *Brain* 2019;142(11):3592-604. doi: 10.1093/brain/awz285.
32. Funkiewiez A, Ardouin C, Caputo E, et al. Long term effects of bilateral subthalamic nucleus stimulation on cognitive function, mood, and behaviour in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2004;75(6):834-9.
33. Dafsari HS, Ray-Chaudhuri K, Mahlstedt P, et al. Beneficial effects of bilateral subthalamic stimulation on alexithymia in Parkinson's disease. *Eur J Neurol* 2019;26(2):222-e17. doi: 10.1111/ene.13773.

34. Yoshida F, Miyagi Y, Kishimoto J, et al. Subthalamic nucleus stimulation does not cause deterioration of preexisting hallucinations in Parkinson's disease patients. *Stereotact Funct Neurosurg* 2009;87(1):45-9. doi: 10.1159/000195719.
35. Zangaglia R, Pacchetti C, Pasotti C, et al. Deep brain stimulation and cognitive functions in Parkinson's disease: A three-year controlled study. *Mov Disord* 2009;24(11):1621-8. doi: 10.1002/mds.22603.
36. Zibetti M, Torre E, Cinquepalmi A, et al. Motor and nonmotor symptom follow-up in parkinsonian patients after deep brain stimulation of the subthalamic nucleus. *Eur Neurol* 2007;58(4):218-23. doi: 10.1159/000107943.
37. Herzog J, Weiss PH, Assmus A, et al. Subthalamic stimulation modulates cortical control of urinary bladder in Parkinson's disease. *Brain* 2006;129(Pt 12):3366-75. doi: 10.1093/brain/awl302.
38. Yamamoto T, Uchiyama T, Asahina M, et al. Urinary symptoms are correlated with quality of life after deep brain stimulation in Parkinson's disease. *Brain Behav* 2018;8(12):e01164. doi: 10.1002/brb3.1164.
39. Castelli L, Perozzo P, Genesia ML, et al. Sexual well being in parkinsonian patients after deep brain stimulation of the subthalamic nucleus. *J Neurol Neurosurg Psychiatry* 2004;75(9):1260-4. doi: 10.1136/jnnp.2003.034579.
40. Kurcova S, Bardon J, Vastik M, et al. Bilateral subthalamic deep brain stimulation initial impact on nonmotor and motor symptoms in Parkinson's disease: An open prospective single institution study. *Medicine (Baltimore)* 2018;97(5):e9750. doi: 10.1097/MD.0000000000009750.
41. Jung YJ, Kim HJ, Jeon BS, et al. An 8-Year Follow-up on the Effect of Subthalamic Nucleus Deep Brain Stimulation on Pain in Parkinson Disease. *JAMA Neurol* 2015;72(5):504-10. doi: 10.1001/jamaneurol.2015.8.
42. Cury RG, Carvalho MJ, Lasteros FJL, et al. Effects of Subthalamic Stimulation on Olfactory Function in Parkinson Disease. *World Neurosurg* 2018;114:e559-e64. doi: 10.1016/j.wneu.2018.03.033.
43. Dafsari HS, Weiss L, Silverdale M, et al. Short-term quality of life after subthalamic stimulation depends on non-motor symptoms in Parkinson's disease. *Brain Stimul* 2018;11(4):867-74. doi: 10.1016/j.brs.2018.02.015.

Table 1 – Baseline characteristics in the original and matched sub-cohort.

	Original cohort (n=151)							Matched sub-cohort (n=76)						
	STN-DBS			MED			p [a]	STN-DBS			MED			p [a]
	n	*M*	*SD*	*n*	*M*	*SD*		*n*	*M*	*SD*	*n*	*M*	*SD*	
Age	67	61.3	8.5	84	65.8	10.4	**.003**	38	62.0	8.9	38	62.5	11.5	.743
Disease duration	67	10.8	4.9	84	7.5	5.0	**<.001**	38	9.9	4.2	38	8.7	5.3	.115
Sex (female/male) [%]	67	25/42	[37.3/62.7]	84	26/58	[31/69]	.412	38	12/26	[31.6/68.4]	38	8/30	[21.1/78.9]	.297
NMSS total score	67	60.6	36.0	83	46.0	25.3	**.010**	38	54.2	34.5	38	46.0	22.1	.436
Cardiovascular	67	1.4	2.3	84	1.0	1.9	.549	38	0.8	1.3	38	1.3	2.3	.489
Sleep/fatigue	67	14.7	10.1	84	9.2	8.6	**<.001**	38	13.0	9.1	38	7.4	6.9	**.007**
Mood/ apathy	67	6.5	11.2	84	6.7	8.4	.215	38	6.3	12.1	38	7.4	8.3	.129
Perceptual problems/ hallucinations	67	1.2	3.1	84	1.1	2.3	.574	38	1.1	3.3	38	1.0	2.3	.621
Attention/ memory	67	5.0	6.2	84	4.6	5.3	.668	38	5.5	5.6	38	5.3	5.1	.900
Gastrointestinal	67	5.8	7.3	84	4.6	5.3	.615	38	4.4	5.5	38	4.9	5.7	.750
Urinary	67	11.5	9.7	84	7.3	6.8	**.018**	38	10.6	9.8	38	6.5	6.3	.122
Sexual function	67	3.4	5.2	84	4.2	5.8	.529	38	3.0	4.7	38	5.0	5.5	.085
Miscellaneous	67	10.9	9.1	83	7.9	6.1	.078	38	9.2	8.7	38	7.3	5.5	.657
PDQ-8 SI	67	32.8	18.5	84	26.2	15.4	**.020**	38	29.1	18.8	38	29.6	14.7	.495
SCOPA-motor examination	67	11.6	5.2	84	10.1	4.6	.077	38	10.8	5.2	38	12.2	4.4	.147
SCOPA-activities of daily living	67	6.9	3.4	84	6.4	3.0	.376	38	6.4	3.8	38	7.2	2.9	.224
SCOPA-motor complications	67	4.8	3.1	84	3.1	1.7	**<.001**	38	3.6	3.0	38	3.4	2.0	.929
LEDD	67	1199.1	587.5	84	778.1	407.1	**<.001**	38	1011.3	518.2	38	913.0	383.6	.540

Legend: Demographic characteristics and outcome parameters at baseline in the original and matched cohorts. Significant results are highlighted in bold font.
[a] Mann-Whitney U test or t test, when parametric test criteria were fulfilled.

Abbreviations: LEDD = levodopa equivalent daily dose; **MED** = standard-of-care medical treatment; **NMSS** = Non-motor Symptom Scale; **SCOPA** = Scales for Outcomes in Parkinson's disease; **STN-DBS** = subthalamic nucleus deep brain stimulation; **PDQ-8 SI** = 8-item Parkinson's Disease Questionnaire Summary Index

Table 2 – Outcomes at baseline and 36-month follow-up in the matched cohort

	STN-DBS						MED						
		Baseline		36-MFU		Baseline vs. 36-MFU [a]		Baseline		36-MFU		Baseline vs. 36-MFU [a]	STN-DBS vs. MED [b]
	n	*M*	*SD*	*M*	*SD*	*p*	*n*	*M*	*SD*	*M*	*SD*	*p*	*p*
NMSS total score	38	54.2	34.5	38.6	24.4	**.018**	38	46.0	22.1	62.8	40.0	.072	**.003**
NMSS domains													
Cardiovascular	38	0.8	1.3	2.4	3.6	**.012**	38	1.3	2.3	1.5	2.8	.833	.117
Sleep/fatigue	38	13.0	9.1	8.6	6.4	**.003**	38	7.4	6.9	12.5	8.8	**.011**	**<.001**
Mood/apathy	38	6.3	12.1	4.2	6.9	.362	38	7.4	8.3	8.1	14.2	.585	.879
Perceptual problems/ hallucinations	38	1.1	3.3	1.0	2.0	.740	38	1.0	2.3	1.6	3.0	.160	.646
Attention/memory	38	5.5	5.6	4.7	5.9	.530	38	5.3	5.1	7.2	8.0	.187	.242
Gastrointestinal	38	4.4	5.5	5.1	4.9	.371	38	4.9	5.7	7.3	7.8	.062	.381
Urinary	38	10.6	9.8	8.0	9.4	**.037**	38	6.5	6.3	12.0	10.4	**.005**	**<.001**
Sexual function	38	3.0	4.7	1.0	2.7	**.031**	38	5.0	5.5	4.3	6.9	.475	.437
Miscellaneous	38	9.2	8.7	3.7	4.5	**<.001**	38	7.3	5.5	8.3	8.8	.703	**.007**
PDQ-8 SI	38	29.1	18.8	23.0	16.4	**.024**	38	29.6	14.7	35.8	18.2	.072	**.004**
SCOPA-motor examination	38	10.8	5.2	8.1	5.0	**.018**	38	12.2	4.4	12.7	5.7	.507	**.026**
SCOPA-activities of daily living	38	6.4	3.8	5.8	3.5	.479	38	7.2	2.9	8.5	4.6	.131	.059
SCOPA-motor complications	38	3.6	3.0	2.1	2.6	**.018**	38	3.4	2.0	4.1	2.4	.072	**<.001**
LEDD (mg)	38	1011.3	518.2	703.9	504.2	**.018**	38	913.0	383.6	981.2	443.1	.550	**.026**

Legend: Outcome parameters at baseline and follow-up for the STN-DBS and MED groups. Multiple comparisons due to multiple outcome parameters were corrected with Benjamini-Hochberg's method. Post-hoc exploratory analyses were performed for NMSS domains. Significant results are highlighted in bold font.
Abbreviations: 36-MFU = 36-month follow-up; **LEDD** = levodopa equivalent daily dose; **MED** = standard-of-care medical treatment; **NMSS** = Non-motor Symptom Scale; **PDQ-8 SI** = 8-item Parkinson's Disease Questionnaire Summary Index; **SCOPA** = Scales for Outcomes in Parkinson's disease; **STN-DBS** = subthalamic nucleus deep brain stimulation.
[a] Wilcoxon signed rank test between baseline and 36-month follow-up to analyze within-group changes of outcome parameters
[b] Mann-Whitney U test were used to analyze between-group differences of change scores between STN-DBS and MED group.

Table 3 – Change scores, Relative changes and Effect sizes for matched cohorts

	Change score		Relative change (%)		Effect size	
	STN-DBS	MED	STN-DBS	MED	Cohen's d	Classification
PDQ-8 SI	6.1	–6.2	21.0	–20.9	0.72	moderate
SCOPA-motor examination	2.8	–0.5	25.7	–4.1	0.67	moderate
SCOPA-activities of daily living	0.6	–1.3	9.4	–18.1	0.56	moderate
SCOPA-motor complications	1.5	–0.7	63.9	–20.6	0.85	large
LEDD	307.4	–68.2	30.4	–7.5	0.82	large
NMSS total score	15.6	–16.8	28.8	–36.5	1.11	large
Cardiovascular	–1.6	–0.2	–200.0	–11.4	0.74	moderate
Sleep/fatigue	4.4	–5.1	33.9	–67.8	1.16	large
Mood/apathy	2.1	–0.7	33.3	–8.9	0.27	small
Perceptual problems/ hallucinations	0.1	–0.6	9.1	–61.0	0.24	small
Attention/memory	0.8	–1.9	14.6	–35.0	0.50	moderate
Gastrointestinal	–0.7	–2.4	–15.9	–50.1	0.30	small
Urinary	2.6	–5.5	24.5	–85.0	0.97	large
Sexual function	2.0	0.6	66.7	12.3	0.25	small
Miscellaneous	5.5	–1	59.8	–14.2	0.88	large

Legend: Change scores and relative changes from baseline to 36-month follow-up in the STN-DBS and MED group. Effect sizes of the between-group comparison STN-DBS vs. MED (Mann-Whitney U test).

Change score $= (mean\ test_{baseline} - mean\ test_{follow-up})$
Relative change $= (mean\ test_{baseline} - mean\ test_{follow-up}) \,/\, mean\ test_{baseline}$ x 100.
Cohen's *d* = (mean pre-post change$_{\text{treatment group}}$ –mean pre-post change$_{\text{control group}}$) / SD pretest$_{\text{pooled groups}}$.
Cohen's *d* can be classified as 'small' ($0.20 \geq d < 0.50$), 'moderate' ($0.50 \geq d < 0.80$) and 'large' ($d \geq 0.80$).
Effect size for the NMSS cardiovascular domain was favorable in the MED group and for all other outcome parameters and NMSS groups in the STN-DBS group.

Abbreviations: LEDD = levodopa equivalent daily dose; **MED =** standard-of-care medical treatment; **NMSS** = Non-motor Symptom Scale; **SCOPA** = Scales for Outcomes in Parkinson's disease; **STN-DBS** = subthalamic nucleus deep brain stimulation; **PDQ-8 SI** = 8-item Parkinson's Disease Questionnaire Summary Index

Table 4 – Spearman correlations

		NMSS total score	Cardio-vascular	Sleep/ fatigue	Mood/ apathy	Perceptual problems/ hallucinations	Attention/ memory	Gastro-intestinal	Urinary	Sexual functions	Miscella-neous
PDQ-8 SI	rho	**.517*** **	.063	**.523*** **	**.256***	.176	**.274***	.079	**.334****	.095	**.377****
	p	<.001	.59	<.001	.026	.129	.016	.497	.003	.413	.001
	n	76	76	76	76	76	76	76	76	76	76

Legend: The relationship between change scores of PDQ-8 SI and NMSS (total and domain scores) was analyzed by Spearman correlations. Significant correlations are marked with stars.

Correlation coefficients for significant results are highlighted in bold.
* $p < 0.050$
** $p < 0.010$
*** $p < 0.001$
Correlation coefficients were 'weak' (rho = 0.200 – 0.399) for mood/apathy, attention/memory, urinary symptoms, sexual functions, and miscellaneous domains, and 'moderate' (rho = 0.400 – 0.599) for the sleep/fatigue domain and the NMSS total score.

Abbreviations: NMSS = Non-motor Symptom Scale; **PDQ-8 SI** = 8-item Parkinson's Disease Questionnaire Summary Index.

Figure 1 – Patient selection

Figure 2 – Non-motor Symptom Scale domains at baseline (blue) and 36-month follow-up (red) for the STN-DBS and MED groups in clustered boxplots (2a) and radar charts (2b)

Legend: Figure 2 illustrates Non-motor Symptom Scale (NMSS) domains at baseline (blue) and 36-MFU (red) for the STN-DBS and MED groups in (A) clustered boxplots and (B) radar charts. Significant within-group changes of NMSS domains from baseline to 36-MFU are highlighted with a black star and significant between-group differences (STN-DBS vs. MED) with a cross. (A) Outliers are illustrated with dots (2-3 standard deviations) and extreme outliers with colored stars (>3 standard deviations). (B) NMSS domain scores are illustrated as percentage of maximum scores. Bigger areas represent more severe NMS impairment.

In the STN-DBS group, NMSS sleep/fatigue, urinary, and miscellaneous domains significantly improved, and the cardiovascular domain significantly worsened from baseline to 36-MFU. In the MED group, NMSS sleep/fatigue and urinary domains significantly worsened from baseline to 36-MFU. Differences in NMSS domain scores were significant for sleep/fatigue, urinary, and miscellaneous domains.

Abbreviations: 36-MFU = 36-month follow up; **MED** = standard-of-care medical treatment; **STN-DBS** = subthalamic nucleus deep brain stimulation